



第一章 大数据与财务管理的基本理论	1
第一节 大数据的发展	1
第二节 企业财务管理概述	15
第二章 中西方管理理论与思想的演进	20
第一节 中国管理思想的演进与镜鉴	20
第二节 西方管理思想的萌芽	29
第三节 中世纪的管理思想和实践	36
第四节 资本主义早期的管理思想	42
第五节 18 世纪后半期到 19 世纪的管理思想	44
第六节 西方现代管理思想的演变	51
第三章 财务管理与大数据的基本关系	60
第一节 企业财务管理的现状	60
第二节 大数据时代下的财务决策的新思维	71
第三节 大数据引发的无边界融合式财务管理	73
第四节 大数据背景下企业的财务决策框架	79
第五节 大数据下的财务管理问题与对策	82
第四章 大数据时代的企业信息管理	86
第一节 企业信息管理概述	86
第二节 大数据与企业信息管理	93
第三节 现代企业信息管理的创新	97

第五章 大数据影响下的财务共享	104
第一节 财务共享的内涵及发展	104
第二节 财务共享中心的建立和运作	108
第三节 企业实现财务共享的影响因素	113
第四节 企业财务共享中心存在的问题及对策	114
第六章 大数据时代的企业专利管理及质量管理	118
第一节 大数据背景下企业专利管理分析	118
第二节 大数据时代的质量管理	122
第七章 企业集团的资本控制、投资控制以及预算控制	133
第一节 企业集团资本控制工作	133
第二节 企业集团投资控制工作	135
第三节 企业集团预算控制工作	138
第八章 预算管理在企业管理中的地位与作用	141
第一节 预算管理控制与战略管理控制	141
第二节 预算管理在企业管理控制系统中的地位	148
第三节 预算管理在企业管理控制系统中的作用	151
第四节 以预算管理为导向的管理控制体系在企业管理中的应用	154
第九章 经济管理体系下的现代企业制度创新研究	158
第一节 现代企业管理制度建设中的问题分析	158
第二节 现代企业管理制度的发展与创新	163
第三节 儒家管理哲学理念与现代企业管理制度的契合	170
第四节 现代企业管理制度的内部控制体系研究	176
第十章 经济管理体系下的现代企业文化创新研究	185
第一节 企业文化在现代企业管理中的作用	185
第二节 企业文化导向的核心竞争力理念	190
第三节 跨文化视域下的企业文化创新	197
第四节 现代企业管理与传统文化的融合创新	201
参考文献	204



第一章 大数据与财务管理的 基本理论

大数据是中国经济新常态下创新驱动的发动机也是产业转型的助推器，从而带动了技术研发体系的创新、管理方式的变革、商业模式创新和产业价值链体系的重构，推动了跨领域、跨行业的数据融合和协同创新。企业可持续发展的关键是获得竞争优势，财务战略作为企业总体战略的核心，基于长期性和系统性的视角，在企业整体目标的引领下，融合了财产的购置、投资、融资以及管理等事项。大数据时代，一方面，海量数据为财务管理和决策带来更有力的支撑；另一方面，也对传统的财务系统、财务人员的素质提出了更高的要求。

第一节 大数据的发展

随着新一代信息技术迅猛的发展与深入应用，数据的数量、规模不断扩大，数据已日益成为土地、资本之后的又一种重要的生产要素，与各个国家和地区争夺的重要资源，谁掌握数据的主动权和主导权，谁就能赢得未来。

一、大数据的定义

互联网的技术不断发展，渗透进我们的工作和生活，加上移动网络、物联网与其他

各种联网设备的出现与普及，一个必然产生的现象就是数据的迅速增长。有90%的数据是互联网出现以后才产生的，它以指数级的速度在我们的生活中不断增加，从海量至于无穷大，世界正被数据淹没。

然而，大数据还没有统一的标准定义。尽管大数据的定义各家歧异，但基本上，大数据领域里的每个人都同意一点：大数据不仅仅是指更多资料。

（一）百度搜索的定义

以容量大、类型多、存取速度快、应用价值高为主要特征的数据集合，最早应用于IT行业，目前正快速发展为对数量巨大、来源分散、格式多样的数据进行采集、存储和关联分析，从中发现新知识、创造新价值、提升新能力的新一代信息技术和服务业态。大数据必须采用分布式架构，对海量数据进行分布式数据挖掘，因此必须依托云计算的分布式处理、分布式数据库和云存储、虚拟化技术。

（二）互联网周刊的定义

互联网周刊的定义为是：“大数据”的概念远远不止大量的数据（TB）以及处理大量数据的技术，而是涵盖了人们在大规模数据的基础上可以做的事情，而这些事情在小规模数据的基础上是无法实现的。

总体来说就是大数据让我们以一种前所未有的方式，通过对海量数据进行分析，获得有巨大价值的产品和服务，或深刻的洞见，最终形成变革之力。

（三）研究机构的定义

研究机构认为：“大数据”是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。从数据的类别上看，“大数据”指的是无法使用传统流程或工具处理或分析的信息。它定义了那些超出正常处理范围和大小、迫使用户采用非传统处理方法的数据集。

二、大数据的内涵

洞见大数据时代发展趋势的数据科学家之一维克·托迈尔·舍恩伯格在他的著作《大数据时代》中指出，大数据不用随机分析法（抽样调查）这样的捷径，而是对所有数据进行分析处理。

大数据研究机构Gartner Group（高德纳咨询公司）给出了这样的定义：“大数据”是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应海量、高增长

率和多样化的信息资产。

《大数据：下一个创新、竞争和生产力的前沿》报告由知名咨询公司麦肯锡所发布，文中阐述了大数据的含义，其主要指代远超普通数据库能完成的数据收集、容纳、掌控和解析技术的数据库。大数据的概念涵盖两个层面：其一是与大数据要求吻合的数据库总量不是一成不变的，它会受到时间和技术的影响而逐步扩大；其二是各个机构对于与大数据要求吻合的数据库总量也存在差异。

综合来看，大数据的内涵至少包含三个要素。

（1）数据性思维。一种可以对海量数据加以分析处理的新模式。人们可以比以往分析更多的数据，甚至是某一事件和现象的全部数据，而不再依赖统计学的随机采样。人们更关心事物的相关性分析，而非因果关系。

（2）数据性资产。拥有海量数据可以洞见众多商机，获得巨大利润。大数据包含的潜在信息正在创造着各种商机，成为重要的商业情报。例如，企业可以通过在汽车上安装各类传感器和行车电脑获取关于汽车零件在汽车行驶时使用情况的实时数据，这种数据可以用于后续的零件设计优化，这会成为企业的巨大竞争优势。

（3）数据技术创新。一种高效率、自动化、平民化的数据处理技术。物联网、云计算等现代数据技术的出现使得大规模运算成为现实，且成本更为低廉。

三、大数据的类型

传统企业数据	包括CRM systems的消费者数据、传统的ERP数据、库存数据以及账目数据等
机器和传感器数据	包括呼叫记录、智能仪表、工业设备传感器、设备日志、交易数据等
社交数据	包括用户行为记录、反馈数据等。如Twitter、Facebook这样的社交媒体平台

四、大数据的特征

虽然到现在为止，全社会对于大数据的认知并没有达成一致。但在学术界内有一种普遍认可的观点，这就是大数据具有“4V”特征。大数据是以体量巨大（Volume）、类型繁多（Variety）、存取速度快（Velocity）、价值密度低（Value）为基本特征的数据集。

（1）体量巨大（Volume）。伴随着各种移动设备、物联网和云计算、云存储等技术的发展，互联网数据已经不局限于传统意义上的网页、论坛。人和物的所有轨迹、状态和行为都可以被记录，因此产生了大量数据。人类社会的数据量在不断刷新一个个新的一级单位，已经从TB（太字节）、PB（拍字节）级别跃升至EB（艾字节）、ZB（泽字节）级别。

(2) 类型繁多 (Variety)。大数据有着多种表现形式。通常人们认为数据只能是整齐地排列成行列样式,而事实上大数据的形式是多样的,而且这种多样性正在不断增加。

在大数据时代,数据格式变得越来越丰富多样,涵盖了文本、音频、图片、视频和模拟信号等不同的类型。Facebook(脸谱网)的照片,移动设备用户点击行为数据,电子地图行为习惯数据,社交网络行为特点和生活习惯,电商交易数据,互联网搜索引擎搜索行为和提问行为,这些数据格式不尽相同。

由于科技进步迅猛,传感器和智能技术有较大提升,各个公司的数据信息越来越趋向于数量大、范围广、难度高。数据不仅涵盖原始的数据关系,而且增加网络、日志等点击量数据,以及搜索引擎、互动媒体、电子信箱、文件、自主和非自发系统的传感器信息等传统、非全结构化数据。

数据采集的规格虽然在相同范围内,但是有所不同。举例而言,北京市区内的智能交通数据采集分析系统的数据是通过路面摄像头、公交车、出租车和各类长途客运、旅游客车、危险品运输车、停车场和租车等取得,还会通过问卷形式和GPS(全球定位系统)获得。整个数据总量涵盖机动车几万辆,公交刷卡每天生成约1900万条数据,手机定位生成约1800万条数据,出租车载客生成约100万条数据,停车场平台系统生成约50万条数据。这些数据虽然产生在相同领域中,而且数量范围和产出效率都满足大数据规格,但数据性质各有不同,产生速度大不相同。

大数据的惊人之处在于其数据库的广泛性。以交通情况为例,它和各种范围的数据都呈现强相关性。通过调查研究,人们能够通过供水公司数据看出晨间洗漱的用水量最大值,再多计算一个误差值,大致能够推测该地区晨间交通高峰期;与此类似,也可以通过用电量看出办公区熄灯时间主要在晚上哪个时间段,再加上误差值,可以大致推测该地区交通晚高峰时间范围。

(3) 存取速度快 (Velocity)。大数据是一种以实时数据处理、实时结果导向为特征的解决方案。大数据的“快速”特征体现在两个方面:一是数据生成非常迅速。数据产出方式不同,有集中式,也有循序渐进式。其中欧洲针对核子探索机构的巨型强子对撞机械运行过程中会在每秒产出大量数据,这属于集中式数据产出。但互联网公司的数据收集是循序渐进的,虽然不是一次性产出,但由于使用者过多,有限时间里也会产出大量数据。Facebook、淘宝、百度和谷歌等,均可以满足每秒数百MB(兆字节)点击率、日志、射频识别数据、GPS位置信息的采集和传输需求,并将这些数据上传到中央系统上。二是数据处理快。举个例子,如果想存储1PB容量的数据,在网速1GB/s(吉字节/秒)前提下,假设计算机存储量能满足全天运行,那么将这么大的数据量完全储存在计算机中需要长达12天的时间。这时大数据的作用便显现出来,它能够通过云技术轻松将计算机12天的工作量在20分钟内处理完毕。

(4) 价值密度低 (Value)。大数据并不直接意味着大价值,实际上大数据的价值之低,也是非日常可以想象的。比如,上亿微信用户中起床第一件事就是发微信的人并不多;一项火热的话题后7万条评论中可能仅有几条具有价值;公司2000名客户代表中也许只有7名客户只喝某个品牌的酒;高炉2000个传感器每分钟上报一个信号,几个月可能仅有5个信号接近临界值;许多产品类别和品牌几乎从来没有在社会化媒体上被提及,样本量可能微乎其微。各类互动平台会被各种外界原因所干扰,发出的议论并不能代表其个人真实的观点。以上这些数据中混杂着大量不相关的、毫无意义的信息,如果再考虑数据采集不及时、数据样本不全面、数据不连续等因素,大数据的含金量可想而知。

五、大数据的技术框架

大数据尝试从海量数据中,通过一定的分布式技术手段,挖掘出有价值的信息,最终提供给用户,进而产生实用价值和商业价值。由于数据本身的多样性以及数据分析需求的多元化,大数据技术体系非常复杂,涉及的组件和模块众多,为了便于读者从顶层框架上对大数据有一个清楚的认识,下面尝试概括大数据技术框架。

在互联网领域,数据无处不在。从数据在信息系统中的生命周期来看,大数据从数据源开始,经过分析、挖掘到最终获得价值,一般需要经过六个主要环节,包括数据收集、数据存储、资源管理与服务协调、计算引擎、数据分析和数据可视化。

(一) 数据收集层

数据收集层由直接跟数据源对接的模块构成,负责将数据源中的数据近实时或实时收集到一起。数据源具有分布式、异构性、多样化及流式产生等特点。

(1) 分布式:数据源通常分布在不同机器或设备上,并通过网络连接在一起。

(2) 异构性:任何能够产生数据的系统均可以称为数据源,比如Web服务器、数据库、传感器、手环、视频摄像头等。

(3) 多样化:数据的格式是多种多样的,既有像用户基本信息这样的关系型数据,也有如图片、音频和视频等非关系型数据。

(4) 流式产生:数据源如同“水龙头”一样,会源源不断地产生“流水”(数据),而数据收集系统应实时或近实时地将数据发送到后端,以便及时对数据进行分析。

由于数据源具有以上特点,将分散的数据源中的数据收集到一起通常是一件十分困难的事情。一个适用于大数据领域的收集系统,一般具备以下几个特点。

(1) 扩展性:能够灵活适配不同的数据源,并能接入大量数据源而不会产生系统瓶颈。

(2) 可靠性：数据在传输过程中不能够丢失（有些应用可容忍少量数据丢失）。

(3) 安全性：对于一些敏感数据，应有机制保证数据收集过程中不会产生安全隐患。

(4) 低延迟：数据源产生的数据量往往非常庞大，收集系统应该能够在较低延迟的前提下将数据传输到后端存储系统中。

为了让后端获取全面的数据，以便进行关联分析和挖掘，通常我们建议将数据收集到一个中央化的存储系统中。

（二）数据存储层

数据存储层主要负责海量结构化与非结构化数据的存储。传统的关系型数据库（比如MySQL）和文件系统（比如Linux文件系统）因在存储容量、扩展性及容错性等方面的限制，很难适应大数据应用场景。

在大数据时代，由于数据收集系统会将各类数据源源不断地发到中央化存储系统中，这对数据存储层的扩展性、容错性及存储模型等有较高要求，总结如下。

(1) 扩展性。在实际应用中，数据量会不断增加，现有集群的存储能力很快将达到上限，此时需要增加新的机器扩充存储能力，这要求存储系统本身具备非常好的线性扩展能力。

(2) 容错性。考虑到成本等因素，大数据系统从最初就假设构建在廉价机器上，这就要求系统本身就有良好的容错机制确保在机器出现故障时不会导致数据丢失。

(3) 存储模型：。由于数据具有多样性，数据存储层应支持多种数据模型，确保结构化和非结构化的数据能够很容易保存下来。

（三）资源管理与服务协调层

随着互联网的高速发展，各类新型应用和服务不断出现。在一个公司内部，既存在运行时间较短的批处理作业，也存在运行时间很长的服务，为了防止不同应用之间相互干扰，传统做法是将每类应用单独部署到独立的服务器上。该方案简单易操作，但存在资源利用率低、运维成本高和数据共享困难等问题。为了解决这些问题，公司开始尝试将所有这些应用部署到一个公共的集群中，让它们共享集群的资源，并对资源进行统一使用，同时采用轻量级隔离方案对各个应用进行隔离，因此便诞生了轻量级弹性资源管理平台，相比于“一种应用一个集群”的模式，引入资源统一管理可以带来众多好处，比如，资源利用率高、运维成本低以及数据共享等。资源利用率高：如果每个应用一个集群，则往往由于应用程序数量和资源需求的不均衡，使得在某段时间内有些应用的集群资源紧张，而另外一些集群资源空闲。共享集群模式通过多种应用共享资源，使得集群中的资源得到充分利用。

运维成本低：如果采用“一个应用一个集群”的模式，则可能需要多个管理员管理这些集群，进而增加运维成本。共享模式通常需要少数管理员即可完成多个框架的统一管理。

数据共享：随着数据量的暴增，跨集群间的数据移动不仅需花费更长的时间，并且硬件成本也会大大增加，而共享集群模式可让多种应用共享数据和硬件资源，这将大大减小数据移动带来的成本。

在构建分布式大数据系统时，会面临很多共同的问题，包括leader选举、服务命名、分布式队列、分布式锁、发布订阅功能等，为了避免重复开发这些功能，通常会构建一个统一的服务协调组件，包含了开发分布式系统过程中通用的功能。

（四）计算引擎层

在实际生产环境中，针对不同的应用场景，我们对数据处理的要求是不同的，有些场景下，只需离线处理数据，对实时性要求不高，但要求系统吞吐率高，典型的应用是搜索引擎构建索引；有些场景下，需对数据进行实时分析，要求每条数据处理延迟尽可能低，典型的应用是广告系统及信用卡欺诈检测。为了解决不同场景下数据处理问题，起初有人尝试构建一个大的系统解决所有类型的数据计算问题，但最终以失败告终。究其原因，主要是因为不同类型的计算任务，其追求的目标是不同的，批处理计算追求的是高吞吐率，而实时计算追求的是低延迟。在现实系统中，系统吞吐率和处理延迟往往是矛盾的两个优化方向：系统吞吐率非常高时，数据处理延迟往往也非常高，基于此，用一个系统完美解决所有类型的计算任务是不现实的。

计算引擎发展到今天，已经朝着“小而美”的方向前进，即针对不同应用场景，单独构建一个计算引擎，每种计算引擎只专注于解决某一类问题，进而形成了多样化的计算引擎。计算引擎层是大数据技术中最活跃的一层，直到今天，仍不断有新的计算引擎被提出。总体上讲，可按照对时间性能的要求，将计算引擎分为三类。

（1）批处理。该类计算引擎对时间要求最低，一般处理时间为分钟到小时级别，甚至天级别，它追求的是高吞吐率，即单位时间内处理的数据量尽可能大，典型的应用有搜索引擎构建索引、批量数据分析等。

（2）交互式处理。该类计算引擎对时间要求比较高，一般要求处理时间为秒级别，这类系统需要跟人进行交互，因此会提供类SQL的语言便于用户使用，典型的应用有数据查询、参数化报表生成等。

（3）实时处理。该类计算引擎对时间要求最高，一般处理延迟在秒级以内，典型的应用有广告系统、舆情监测等。

（五）数据分析层

数据分析层直接跟用户应用程序对接，为其提供易用的数据处理工具。为了让用户分析数据更加容易，计算引擎会提供多样化的工具，包括应用程序API、类SQL查询语言、数据挖掘SDK等。在解决实际问题时，数据科学家往往需根据应用的特点，从数据分析层选择合适的工具，大部分情况下，可能会结合使用多种工具，典型的使用模式是：首先使用批处理框架对原始海量数据进行分析，产生较小规模的数据集，在此基础上，再使用交互式处理工具对该数据集进行快速查询，获取最终结果。

（六）数据可视化层

数据可视化技术指的是运用计算机图形学和图像处理技术，将数据转换为图形或图像在屏幕上显示出来，并进行交互处理的理论、方法和技术。它涉及计算机图形学、图像处理、计算机辅助设计、计算机视觉及人机交互技术等多个领域。

数据可视化层是直接面向用户展示结果的一层，由于该层直接对接用户，是展示大数据价值的“门户”，因此数据可视化是极具意义的。考虑到大数据具有容量大、结构复杂和维度多等特点，对大数据进行可视化是极具挑战性的。下面我们举例说明发展可视技术的意义及挑战。在医学领域，为了认识人体内部结构，美国国家医学图书馆于1989年开始实施可视化人体计划（VHP），并委托科罗拉多大学医学院建立了一男一女的全部解剖结构数据库。他们分别将男女不同性别的两具尸体从头到脚做CT扫描和核磁共振扫描（男的间距1毫米，共1878个断面；女的间距0.33毫米，共5189个断面），然后将尸体填充蓝色乳胶并裹以明胶后冰冻至零下80摄氏度，再以同样的间距对尸体作组织切片的数码相机摄影，分辨率为2048x 1216，最终所得数据共56GB（男13GB，女43GB）。全球用户可以在美国国家医学图书馆允许的情况下获得该数据并用于教学和科学研究。VHP数据集的出现标志着计算机三维重构图像和虚拟现实技术进入了医学领域，从而大大促进了医学的发展和普及。

六、大数据基础技术

大数据处理流程主要是指从海量数据中获取需要的信息数据并进行加工分析得到有益知识的输出过程。大数据本质也是数据，其处理流程的关键技术依然逃脱不了大数据存储和管理及大数据检索使用（包括数据挖掘和智能分析）。围绕大数据，一批新兴的数据存储、数据挖掘、数据处理与分析技术不断涌现，让我们处理海量数据更加容易、更加便宜和迅速的同时，也成为企业业务经营的好助手，甚至可以改变许多行业的经营方式。因

此,一般将大数据处理流程概括为以下几个步骤:数据采集和清洗、数据存储、数据挖掘以及数据呈现。

(一) 数据采集和清洗

数据采集又称“数据获取”,是指从传感器与其他待测设备等模拟和数字被测单元中自动采集信息的过程。在互联网行业快速发展的今天,数据采集已经被广泛应用于互联网及分布式领域,数据采集领域已经发生了重要的变化。

随着互联网的普及,每个人每天都在网络上产生无数的数据,这些数据构成了大数据的基础。想要收集这些数据,能使用的主要工具有传感器、网络爬虫、移动基站和使用者自身产生的信息。

(1) 传感器。传感器是一种检测装置,能感受到被测量的信息,并能将感受到的信息,按一定规律变换成电信号或其他所需形式的信息输出。分布在生活中的传感器产生了海量的数据来源,如监控大型强子对撞机或四发动机大型喷气式客机需要成千上万的传感器通道,产生数百TB的数据。

(2) 网络爬虫。网络爬虫是按照一定的规则,自动抓取互联网网页信息的程序或者脚本。

(3) 无源光纤网络。日常的通信过程产生海量信息。

(4) 使用者自身产生的信息。微信、微博、邮件等渗入大家生活,而拥有庞大用户群的微信、微博或者淘宝网上都留下了巨量的信息,这些信息便是海量数据的来源之一。

(二) 数据库和数据存储

数据库要主要解决以下问题:如何存储数据、如何找到数据、如何定义数据以及如何管理数据。数据库的基本组成包括存储数据的实体、管理数据的方式和方法。

数据的存储主要关乎数据分类、数据索引、数据属性(直白点就是仓储管理)。

数据存储技术主要包括裸数据和文件系统。

具体定义如下:

(1) 数据。数据是指描述事物的符号记录,可用文字、图形等多种形式表示,经数字化处理后可存入计算机。

(2) 数据库(DB)。数据库是按一定的数据模型组织、描述和存储在计算机内的.有组织的、可共享的数据集合。

(3) 数据库管理系统(DBMS)。数据库管理系统是位于用户和操作系统之间的一层数据管理软件。主要功能包括:数据定义功能:DBMS提供DDL,用户通过它定义数据对

象。数据操纵功能：DBMS提供DML，用户通过它实现对数据库的查询、插入、删除和修改等操作。

1.信息世界

实体：客观存在并可相互区分的事物。

实体集：性质相同的同类实体的集合。

属性：实体具有的某一特性。

实体标识符：能使一个实体与其他实体区分开来的一个或一组属性

2.数据世界

记录 $\longleftrightarrow$ 实体（抽象表示）

文件 $\longleftrightarrow$ 实体集

字段或数据项 $\longleftrightarrow$ 属性

关键字 $\longleftrightarrow$ 实体标识符，唯一地标识一个记录，又称码、键。

3.实体——联系方法（ Entity - Relationship Approach ）用E-R图（ Entity - Relationship Diagram ）描述：

（1）实体型，用长方形表示；联系，用菱形表示；属性，用椭圆形表示。

（2）框内写上相应的名称。

（3）用无向边连接。

4.数据模型

数据模型是对现实世界进行抽象的工具，它按计算机系统的观点对数据建模，用于提供数据库系统中信息表示和操作手段的形式框架，主要用于DBMS的实现，是数据库系统的核心和基础。

数据定义语言（DDL）。

数据操作语言（DML）。

数据查询语言（DQL）。

数据控制语言（DCL）。

（1）信息世界。

实体：客观存在并可相互区分的事物。

实体集：性质相同的同类实体的集合。

属性：实体具有的某一特性。

实体标识符：能将一个实体与其他实体区分开来的一个或一组属性。

（2）数据世界。

记录 $\longleftrightarrow$ 实体（抽象表示）

文件 $\longleftrightarrow$ 实体集

字段或数据项 $\longleftrightarrow$ 属性

关键字 $\longleftrightarrow$ 实体标识符，唯一地标识一个记录，又称码、键。

关系：对应一张表，每表起一个名称即关系名。

元组：表中的一行。

属性：表中一列，每列起一个名称即属性名。

主码：唯一确定一个元组的属性组。

域：属性的取值范围。

数据库：一个或多个数据库。

数据库的四要素：用户数据、元数据、索引和应用元数据。

（三）数据挖掘和数据分析

在得到所需要的数据后，如何从这些海量数据中得到有效的信息，去粗取精得到用户需要的信息，这便是数据挖掘的过程，数据挖掘是一种决策支持过程。它主要基于人工智能统计学、数据库、可视化技术等，高度自动化地分析企业的的海量数据，做出归纳性地推理，从中挖掘出潜在的模式，帮助决策者调整市场策略，减少风险，做出正确的决策。大数据分析中，主要使用的数据挖掘手段有以下几种。

1. 聚类分析：如基于历史的MBR分析、遗传算法。

聚类分析是把一组数据按照相似性和差异性分为几个类别，其目的是使得属于同一类别数据间的相似性尽可能大，不同类别数据间的相似性尽可能小。它可以应用到客户群体的分类、客户背景分析、客户购买趋势预测、市场的细分等方面。

基于历史的MBR分析方法最主要的概念就是用已知案例来预测未来案例的一些属性，通常寻最相似的案例来做比较。MBR分析中有两个主要因素：距离函数与结合函数。距离函数用于找出最相似的案例，结合函数则将相似案例的属性结合起来，以供预测之用。MBR分析方法的优点是容许各种形态的数据。

遗传算法是一种基于生物自然选择与遗传机理的随机搜索算法，是一种仿生全局优化方法。遗传算法具有的隐含并行性、易于和其他模型结合等性质使得它在数据挖掘中被加以应用。Sunil已成功地开发了一个基于遗传算法的数据挖掘工具，利用该工具对两个飞机失事的真实数据库进行了数据挖掘实验，结果表明遗传算法是进行数据挖掘的有效方法之一。

2. 关联分析：如购物车分析。

关联分析就是从大量数据中发现项集之间有趣的关联和相关联系。关联分析的一个典型例子是购物车分析。购物车分析就是通过发现顾客放入购物车中的不同商品之间的联系，分析顾客的购买习惯，主要目的在于找出什么样的东西应该放在一起。根据顾客的购

买行为找出相关的联想规则，这种规则的挖掘发现可以帮助零售商制定营销策略，零售商可据此分析改变置物架上的商品排列或设计吸引客户的商业套餐等，如发生在美国沃尔玛连锁超市的真实案例——“尿布与啤酒”。

3.分类分析：如决策树、判别分析。

分类分析区别于聚类分析，分类分析是有监督的学习。其中决策树是数据挖掘分类算法的一个重要方法。在各种分类算法中，决策树是最直观的一种。它通过将大量数据有目的地分类，从中找到一些有价值的、潜在的信息。它的主要优点是描述简单，分类速度快，特别适合大规模的数据处理。判别分析是根据表明事物特点的变量值和它们所属的类，求出判别函数，根据判别函数对未知所属类别的事物进行分类的一种分析方法，其核心是考察类别之间的差异。

（四）大数据分析 with R 语言

Hadoop的分布式数据处理模式，让原来不可能的TB、PB级数据量计算成为可能，而R语言的强大之处在于统计分析，但是对于大数据的处理，我们只能通过抽样计算。由Revolution Analytics公司发起的一个开源项目RHadoop，对Hadoop集群上存储在Hadoop分布式文件系统的数据进行本地R分析，并对这些计算的结果进行整合，类似Map Reduce对非结构化数据的操作，这一举措使得R语言与Hadoop结合在一起，可以为用户处理千兆级的数据集。同时，也解决了标准R语言软件包常遇到的内存屏障问题。由此可以看出，这两种技术的结合既是产业界必然的导向，也是产业界和学术界的交集，更为交叉学科的人才提供了无限广阔的想象空间。那么，Hadoop与R到底是如何结合的呢？主要从以下几个方面分析。

（1）RHadoop。RHadoop包含三个R包（Rmr、Rhdfs、Rhbase），分别对应Hadoop系统架构中的MapReduce、HDFS、HBase三个部分。

（2）RHive。RHive是一款通过R语言直接访问Hive的工具包，是由NexR公司研发的。通过使用RHive可以在R环境中写HQL（HiveQL），将R的对象传入Hive中，在Hive中进行计算。在RHive中，小数据集在R中执行，大数据集在Hive中运行。

（3）Hadoop调用R。R可以调用Hadoop，同样地，打通Java与R的连接通道，让Hadoop也可以调用R。

七、大数据的发展趋势

大数据领域的发展激动人心，它具有彻底改变企业经营模式的能力。因此，我们很容易只看到大数据的优点，而忽略它的缺陷。大数据的使用固然重要，但并不是无限制的，

大数据的道德问题也很重要，尤其在商业领域，道德问题也非常重要，企业声誉可谓是成功的关键。有了社交媒体，丑闻眨眼间就能传播到世界各地，多年来树立的企业形象将在瞬间毁于一旦。

如果你听说过一些预测模型，以及企业利用大数据和分析技术从事的业务，就了解到什么是真正的危险，那就是隐私要给可能性让步。很明显，这就涉及透明化问题，企业需要认真管理透明化问题。

在大数据领域，还需要解决重要的道德和伦理问题。大数据和“淘金热”有点相似——没有法律限制，又具有无限的机遇，人们甘心冒早期风险。但相关法律很快就能健全，越来越多的人会为此感到不舒服。例如，企业收集了哪些信息？用来做什么？现在可能发生什么？

（一）迎接大数据的反作用力

可以预见，大数据会有反作用力。在参加的每个研讨会或主旨演讲中，人们经常会为数据收集的发展水平感到震惊，他们并没有意识到，自己很轻易地就会同意别人使用自己的信息，几乎是不假思索地。几乎没有人清楚脸书（Facebook）做过多少分析，可以根据点赞和更新对你有多深入的了解。和其他伟大的创新一样，大数据也是一把双刃剑，仅仅是人脸识别软件带来的可能性就足以让我们感到毛骨悚然。人脸识别软件既可以用来预防犯罪，打击恐怖主义，也可以为实现商业目的监视普通人。在一些官方上并不存在的地方，或大型企业的地下室中，大量的数据可以被获取，正在被很多先进技术孕育着，而大多数人并没有意识到这一点，这是我们面临的最大挑战之一。

我们并不清楚别人收集了哪些数据信息，尽管企业或应用程序已经在合同条款中明确列出，但大多数人都不會阅读这些条款，即使阅读了，也不明白其中的意思，或者弄不明白同意这些条款意味着什么。

举个简单的例子：在伦敦咖啡馆，顾客需要同意相关条件和协议，才能免费使用无线网络。试验中有一条条款规定用户需要“将自己的第一个孩子交给公司”，才能使用免费网络，试验中，有几个人居然欣然接受。

很多小型企业使用谷歌的电子邮件服务，谷歌邮箱既免费又可靠，而且存储容量大。但谷歌认为，顾客在使用他们提供的免费服务时，就没有什么隐私了，但我们对这些又了解多少呢？简单来说，谷歌认为它可以阅读和分析谷歌邮箱用户发送或接收的任何邮件内容。美国联邦法院在案情提要中提到这个事实，并用来起诉谷歌。由于谷歌浏览用户邮件，被指控违反美国联邦和州法律，在辩护中发表了以下声明（最近由消费者保护组织曝光）：如果给同事发邮件，收件人的助理打开邮件，发件人不会为此感到惊讶。同样，发送电子邮件时，在收件过程中，收件人的ECS提供商打开邮件也再正常不过了。

因此,从本质上来看,如果我们注册使用谷歌邮箱,就意味着放弃了所有的隐私权。谷歌通过使用文本分析,得到有用信息,进而优化广告定位。据本人猜测,在G-mail的4亿多用户中,95%的用户现在还未意识到这一点。谷歌邮箱只是其中之一,脸书(Facebook)因一直修改隐私政策和隐私设置而“闻名”或臭名昭著。

任何人都明白企业需要盈利,提供像谷歌邮箱这类免费服务对一些人来说就是足够的回报了。很多人并不关心隐私权。但如果我们想要平安蹊过浑水,必须加大透明度,明确收集了哪些数据,如何使用或将会如何使用这些数据。

关于数据和法律,很多法律体系都在完善。例如,斯堪的纳维亚国家的数据保护法要比英国和美国严厉得多。据预测,英国和美国将实施新的立法,加强保护数据和个人隐私。企业需要在他们收集的数据内容和原因方面更加坦率,而消费者对于他们的数据也将拥有更自由的选择。

(二) 透明化和道德化

现在,很多数据收集行为并不道德。例如,脸书将很多与数据有关的条款放在50页的用户协议中,而用户根本不会阅读用户协议。企业向顾客解释收集了哪些数据以及将如何使用这些数据是非常重要的。如果一家公司并不是非常道德的企业,企业目标只是收集更多的数据,而不关心别人的想法,那么这种做法在短期内可能有效,能够收集大量数据,但长期看来,对企业并没有什么好处。

长期来看,公司道德水准越高,收集的数据价值越大。如果不能坦率地表明收集和存储了顾客的哪些信息,那么企业就有失去这些数据的风险,企业名声也可能受损。如果顾客清楚企业收集了什么数据,以及如何使用这些数据,一般来说顾客都很乐意让企业使用这些数据,任何人都不喜欢发现自己上当受骗。

例如,消费者买了一块崭新的苹果手表,很乐意让苹果公司收集相关数据(比如睡眠模式或每天走多少步),因为公司很清楚他们追踪什么数据,而且公司收集的这些数据有利于消费者找到更健康的生活方式,所以消费者能从中获得益处。但是,假设苹果公司将数据出售给保险公司,保险公司利用这些数据修改保险金,那么消费者是不会乐意的。

遵循以下几条数据透明化建议,消费者会更乐意买单。

第一,如果需要收集消费者或员工的信息,要坦率直言。

第二,解释收集这些信息的原因(例如,能提供更好的服务)。

第三,不要将这些信息隐藏在冗长的用户协议或条款中,消费者根本不会阅读这些协议。相关信息简单明了,并且置于醒目的位置。消费者在网购注册详细信息时,用几句话带过,这样就非常好。

第四,消费者提供宝贵的数据应得到相应的回报(例如,消费者参加调查可以享受折

扣，或者说如果允许记录消费者的数据，可以简化其购物流程）。

第五，消费者拥有选择退出的权利。尽管这意味着他们将不再使用公司提供的服务或部分服务，但最好还是让消费者拥有退出的权利。

第六，尽可能使用聚合数据，聚合数据与具体的个人没有联系。例如，将信息提供给对流行趋势和热门话题感兴趣的第三方，而流行趋势和热门话题与个人没有关系，只是一个群体追求的整体趋势。

（三）确保有附加价值

当企业收集别人的数据时，诚实非常重要，如果能给消费者提供附加价值——让消费者觉得提供这些数据是值得的，也是个不错的主意。

例如，消费者购买了一台最新款的智能电视，可以在这台电视上进行设置，而且能够使用内置摄像头。这台智能电视能够对家里的孩子进行面部识别，对他们观看的电视节目进行限制。消费者大多不介意电视厂家了解我观看了哪些节目，什么时候观看，观看时间的长短，因为这能帮助家长防止孩子观看不适合他们的节目。但是，当人们得知该厂家在搜集人们观看电视的数据后，就陷入了麻烦。那么，提高透明度和增加附加价值可以解决这类问题。

企业应该通过提供质量更好或价格更便宜的产品和服务，为分享个人信息的人带来实惠。始终寻求提供附加价值，这样一来，提供数据的消费者、员工以及其他股东都会认为这种交换是平等的，从而实现多赢。

如果人们能从中获益，那么人们会很乐意为企业提供数据——尤其是当企业能去掉信息的个人标签时。如果能向大家证明企业会合法使用这些信息，人们会做出积极响应。从根本上说，这么做有利于提高数据的长期价值。如果让消费者感觉企业侵犯了自己的隐私，企业会因此失去大批的顾客，在这种情况下，利用数据更深入了解顾客就百害而无一利了。

第二节 企业财务管理概述

一、企业财务管理基本介绍

企业从事商品生产经营活动需要筹集一定数额的资金，并将其直接或间接的投入商品

生产经营活动，当企业售出商品取得收入以后还应对实现的收益进行分配，企业的这一活动称为企业的财务活动，简称企业财务。实际上，企业财务活动是很复杂的，它涉及社会的各个领域。为了保证企业财务活动的正常进行和发挥其在生产经营活动中的作用，企业必须加强财务管理。企业财务管理是指对企业财务活动所进行的管理，它是企业的一项综合性价值管理，在企业管理中占有重要地位。要全面、深入了解企业财务管理的含义，必须进一步掌握企业财务活动的本质。

企业财务管理的内容，包括固定资金管理、流动资金管理、销售收入和利润的管理、专用基金管理等。有另一种观点认为，企业财务管理的主要内容为筹资管理、投资管理、营运资金管理以及利润分配管理。在一个企业中，财务管理最重要的职能就是对生产经营活动发挥指导作用。财务部门最注重数据资料研究，财务人员得到数据后，利用自身专业的敏感度，能够迅速分析业务的变动情况，从而提升对经济活动的洞察能力。

二、企业财务活动

企业财务管理是企业管理的一部分，它是根据财经法规制度，按照财务管理原则组织企业财务活动、处理财务关系的一项经济管理工作。它一般包括筹资管理、投资管理、营运资金管理和利润分配管理四项主要内容。

（一）筹资管理

筹资是指企业为了满足投资和用资需要，筹措和集中所需资金的过程，筹集的资金是资金运动的起点。总体来说，企业可以从两方面筹资并形成两种性质的资金来源：一是企业权益资金；二是企业债务资金。

（二）投资管理

投资是指以收回本金并取得收益为目的而发生的现金流出。企业投资可分为广义的投资和狭义的投资两种。广义投资包括对外投资，如购买其他公司股票、债券或与其他企业联营和对内投资。狭义的投资仅指对外投资，如购置固定资产、无形资产、流动资产等。

（三）营运资金管理

在企业的日常生产经营活动中，会发生一系列的资金收付行为，如销售商品、提供劳务、采购材料、支付工资等。这种因企业经营而引起的财务活动，称为资金的营运活动。在一定时期内，营运资金周转速度越快，资金的利用效率就越高，企业就可能生产出更多的产品，取得更多的收入，获取更多的利润。

（四）利润分配管理

利润分配过程实际是利润在国家税收、企业留存收益和投资者股利之间的划分。所以，广义的分配是指对企业各种收入进行分割和分派的行为；狭义的分配仅指对企业净利润的分配。

三、企业财务关系

企业在处理各项财务活动的过程中，会与多个利益主体发生经济利益关系，称为财务关系，一般包括以下七种关系。

（一）企业与投资者之间的财务关系

企业与投资者之间的财务关系主要是指企业的投资者向企业投入资金，企业向其投资者支付投资报酬所形成的经济关系，主要体现为经营权与所有权的关系。

（二）企业与债权人之间的财务关系

企业与债权人之间的财务关系主要是指企业向债权人借入资金，并按合同的规定偿还本金和支付利息所形成的经济关系，主要体现为债务和债权的关系。

（三）企业与受资者之间的财务关系

企业与受资者之间的财务关系主要是指企业以购买股票或直接投资的形式对其他企业投资所形成的经济关系，主要体现为投资和受资的关系。

（四）企业与债务人之间的财务关系

企业与债务人之间的财务关系主要是指企业将其资金以购买债券、提供借款或商业信用等形式出借给其他单位所形成的经济关系，主要体现为债权和债务的关系。

（五）企业与政府之间的财务关系

企业与政府之间的财务关系是指政府作为社会管理者，以提供公共行政管理和收取各种税款而与企业形成的经济关系，主要体现为依法纳税和依法征税的关系。

（六）企业内部各单位之间的财务关系

企业内部各单位之间的财务关系主要是指企业内部各单位之间在生产经营各环节中互相提供产品或劳务所形成的经济关系，主要体现为企业与内部各单位之间的关系。

（七）企业与职工之间的财务关系

企业与职工之间的财务关系主要是指企业向职工支付劳动报酬过程中所形成的经济利益关系，主要体现为劳动成果的分配关系。

四、财务管理的目标品

财务管理目标（Goals of Financial Management），是指企业进行财务活动所要达到的根本目的。一般而言，最具有代表性的观点包括以下三种。

（一）利润最大化观点

亚当·斯密关于“经济人”假说的利润最大化目标是经济学界的传统观点，时至今日这种观点在理论界与实务界仍有较大影响。这种观点以追逐利润最大化作为财务管理的目标，但这种观点存在以下不足。

（1）没有考虑利润取得的时间。例如，某企业今年获得利润100万元与一年后获得利润100万元，哪一个利润的价值更大呢？如果单从利润额上判断，就会得出经济效益相同的错误结论。因为货币的时间价值是客观存在的，现在100万元的价值要大于一年后的100万元的价值。

（2）没有反映创造的利润与投入的资本之间的关系。例如，两个条件基本相同的企业同样获得100万元利润，一个企业投入500万元资金，另一个企业投入800万元资金，哪一个更符合企业财务目标呢？答案肯定是前者，因为两个企业所获得的利润相同，但投入额却不同，如果不考虑投入产出的比率，就不能做出正确的财务判断。

（3）没有考虑获取利润和所承担风险的关系。例如，两个企业同样投入1000万元，本年同样获得利润300万元，一个企业的利润全部实现了现金收入，另一个企业的利润全部为应收账款，有可能发生坏账损失。哪一个企业的财务管理更符合企业的财务目标？答案是显而易见的。因此，如果不考虑风险，是难以做出正确的财务决策的。

（二）权益资本净利润率最大化或每股收益最大化

权益资本净利润率是利润额与资本额的比率，每股收益是利润额与普通股股数的比值，这里利润额是净利润。这个目标的优点是把企业实现的利润额和股本投入的资本进行对比，能够说明企业的盈利水平，可以在不同资本规模的企业或同一企业不同期间进行比较，揭示其盈利水平的差异。但该指标没有考虑资金时间价值和风险因素，也不能避免企业的短期行为。